

# Assessment 2.0

R J Elliott

Scottish Qualifications Authority, Glasgow, United Kingdom

**Abstract**—This paper considers current assessment practice, looks at the impact of the Internet on today’s learners, and explores ways of modernizing assessment to narrow the gap between the everyday lives of students and the assessment practices that we impose on them.

**Index Terms**—Assessment, Web 2.0, modernize, learning.

## I. ASSESSMENT 1.0

At its most basic level, assessment is the process of generating evidence of student learning and then making a judgment about that evidence. Current assessment practice provides evidence in the form of examination scripts, essays and other artifacts.

For the purposes of this paper, ‘Assessment 1.0’ can be thought of as assessment practice from the beginning of the 20<sup>th</sup> century until today. Throughout this period, assessment exhibited the following characteristics:

- paper-based
- classroom-based
- formalized (in terms of organization and administration)
- synchronized (in terms of time and place)
- controlled (in terms of contents and marking).

These characteristics have changed little during this period; a school master from 1907 would feel at home in an examination hall in 2007.

This assessment system has served us well. The highly centralized, top-down, industrialized system matched the kind of society that existed throughout most of the 20<sup>th</sup> century. Its stability has engendered widespread public confidence in the examination system in the UK (QCA, 2006) and maintained national qualifications as the primary means of entry to employment and Higher Education. The system is also widely understood by its users (students, parents, teachers, university admissions staff, employers and politicians) being relatively unchanged from generation to generation.

A more up-to-date form of assessment has emerged in the last decade, which involves the use of computers in the assessment process. ‘E-assessment’ embraces ‘e-testing’ (a form of on-screen testing of knowledge) and ‘e-portfolios’ (a digital repository of assessment evidence normally used to assess practical skills). A number of commercial products have emerged such as *Questionmark* (e-testing) and *Pebblepad* (e-portfolio). These dedicated systems provided specialized facilities to support online testing or online portfolio building.

In recent years, traditional assessment has been the subject of criticism. The current system is struggling to cope with the demands being placed on it. It was designed to filter students by ability for the purpose of employment or

university selection – not mass accreditation of student achievement.

Because of its bureaucratic nature, it’s expensive to run and doesn’t scale well. Awarding bodies’ costs are rising and these are being passed onto schools and colleges, which complain about the rising burden of examination fees. It’s also inflexible, organized around annual examination ‘diets’.

In addition to these practical problems, there are educational and political concerns. Some educationalists claim that the current assessment system encourages surface learning and “teaching to the test”. Instead of instilling genuine problem solving skills, it fosters memorization. Examination papers that appear to pose ‘deep’ questions are answered by rote memory – memories that are acquired by students under pressure from parents who want to see their children gain qualifications, and drilled by teachers who are seeking to meet targets. Employers complain that, in spite of rising achievement (DfES, 2006), young people are not gaining the skills that are needed in the modern workplace – skills such as problem solving, collaboration, innovation and creativity. Teachers complain about the rising burden of time spent carrying-out and marking assessments, which reduces the time available for “real teaching”. Students complain that the only time that they are required to undertake extended hand-writing is during an examination.

These criticisms are not confined to paper-based assessment. E-testing has been criticized for crudely imitating traditional assessment; vendors of computer-based testing systems boast about their systems’ “faithful reproduction of the paper experience”. These systems typically support a limited number of question types (almost always selected response questions) and, at best, crude simulations of traditional tasks. Some high profile simulations have proven to be unreliable (QCA, 2006), in spite of a great deal of expenditure, leading some commentators to conclude that simulations have inherent reliability problems – problems not faced by “real” assessments. Most contemporary e-portfolio systems, likewise, set-out to mirror the existing curriculum, effectively little more than online storage for students’ work, with a highly content-focused (rather than student-centered) approach to assessment.

These criticisms of e-assessment mirror the criticisms of virtual learning environments (VLEs) – that they simply seek to mimic traditional classroom practice; the “primacy of pedagogy” as Cousin (2004) described it: “VLE environments (*sic*) tend to be skewed towards the simulation of the classroom, lecture hall, tutor’s office and the student common room.” Similarly, most contemporary e-assessment systems are skewed towards the simulation of the class test and the examination hall; or, to paraphrase Cousin, they re-enforce the “tyranny of testing” rather

than seek original and authentic ways of assessing student learning.

Both paper-based and computer-based assessments are perceived by students as something external to them; something over which they have no control; something that is 'done' to them. And the assessment instrument itself is considered contrived, just a hurdle to be jumped, not part of their learning. Or, worse, it is perceived as the sole purpose of their learning, with all their efforts going into passing the test rather than the acquisition of new knowledge and skills. Both forms of assessment also tend to focus on factual knowledge rather than deeper levels of learning, which is difficult to assess using current systems. Yet, factual knowledge is valueless in a culture where Google and Wikipedia is a mouse-click away.

Assessment 1.0 (and 1.5) is also intensely individualistic. Assessment activities are done alone, competition is encouraged, and collaboration (or "cheating" in the lexicon of Assessment 1.0) is prohibited. Assessment 1.5 inherited Assessment 1.0's obsession with security, with products proclaiming that "students are completely disconnected from the network", with "no access to their familiar desktop tools". Not ideal preparation for the 'networked information economy'.

While the familiarity of VLEs has encouraged reticent teachers to experiment with them, it has been claimed (JISC 2006) that their use can actually reduce innovation in the classroom by atrophying classroom practice into traditional (classroom-based) and new (VLE-based), rather than encourage the full potential of e-learning to be explored and applied in the classroom. The use of e-assessment systems might likewise hold back progress in assessment by similarly constraining practice to traditional (paper-based) assessment and the limited form of computer-based assessment made possible by current systems.

The process of moving from paper-based systems to e-testing systems often sheds light on the dark corners of traditional assessment. Existing questions are often found wanting when benchmarked against the standards required for an item bank: rubrics are frequently found to be incomplete (with the missing bits "understood" by the markers), the purpose of questions (what they are trying to assess) is often vague or not known, and there is rarely a consistent approach to their categorization (such as their level of difficulty).

## II. WEB 2.0

Meanwhile, the Internet is evolving. 'Web 2.0' is the name given to the current state of development. Anderson (2006) describes "six big ideas behind Web 2.0". These are:

- user-generated content
- the power of the crowd
- data on an epic scale
- architecture of participation
- network effects
- openness.

For the purposes of this paper, four of these ideas are of particular relevance.

**User-generated content** refers to the ease of creating content. Web services such as Bebo, Wordpress and YouTube have made it easy to create content – and more and

more young people are doing so, with social networking sites becoming a significant part of contemporary culture.

The **power of the crowd** refers to the collective intelligence that can be harnessed from large groups of people. The basic premise is that, subject to certain conditions, a large group of knowledgeable (but non-expert) users can make better decisions than any individual expert. Web services such as Digg and Wikipedia are cited as examples of this collective intelligence.

**Architecture of participation** is based on the twin ideas that Web services must be easy to use (thereby encouraging participation) and organized in such a way as to improve as more people use them. Google Search is a good example since it is very straight-forward to use and its search algorithms learn from the results of previous searches. An aspect of ease-of-use is the idea that not only is new content easy to create but it should be easily created from pre-existing content or easily combined with the contents of other web services ("mash-ups").

**Openness** not only refers to the use of open source software for many Web 2.0 services but also the philosophy of the free sharing of information and resources among users, making it relatively straight-forward to capture and share information or resources, such as embedding a YouTube video in a blog. The generous copyright terms of Creative Commons licenses illustrate this philosophy.

## III. THE NATURE OF CONTEMPORARY LEARNERS

It is in this environment that today's students are living and learning. In *Digital Natives, Digital Immigrants*, Prensky (2001) argues that there is a fundamental distinction to be made between today's learners and those of the past due to "the arrival and rapid dissemination of digital technology... an event which changes things so fundamentally that there is absolutely no going back". He labeled these new learners "digital natives" and contrasted them with "digital immigrants": "The single biggest problem facing education today is that our digital immigrant instructors, who speak an outdated language (that of the pre-digital age), are struggling to teach a population that speaks an entirely new language".

Today's learners are also known by other names. Diana Oblinger (2003) of Microsoft calls them the 'Millennial generation': "Millennials exhibit distinct learning styles. For example, their learning preferences tend toward teamwork, experiential activities, structure and the use of technology. Their strengths include multitasking, goal orientation, positive attitudes, and a collaborative style". From the student's perspective, 'Net-Geners' are "academically driven... we refuse to accept elders' speeches or sermons at face value... our technological savvy makes us smarter, easily adaptable, and more likely to employ technology to solve problems" (Windham, 2005).

A common set of characteristics emerges from the literature on the 'digital native' with respect to their learning styles. These are:

- skilled use of tools
- active learning rather than passive receiving of knowledge
- authentic learning experiences rather than contrived tasks
- task (not process) oriented

- just in time learning
- search not memorize
- doesn't know answer but knows where to find it
- Google not libraries
- collaborate not compete.

When tasked with an assignment, a young person is likely to look-up Wikipedia, search for relevant information on Google, seek help from their friends via Hotmail or MSN, finally pulling together the resulting information into a coherent document using a range of web-based and desktop applications. Unless, of course, the assignment is the same as last year's, in which case a simple e-mail to a friend (or someone else in their extended social network) requesting last year's answer will be sufficient for these goal-oriented learners.

The above scenario sidelines the formal teaching and learning that the student is meant to follow. There is a growing disconnection between the lives of students inside and outside of the classroom. "Schools should not expect students to leave the 21<sup>st</sup> century in the cloakroom; for example, many schools do not allow e-mail, instant messaging, mobile phones or blogging" (Owen *et al* 2006). And the list of prohibited technologies is growing. Twist and Withers (2006) describe the ways in which young people really learn as the "hidden curriculum" – the "informal digital spaces", such as Facebook and MSN, which students routinely use for social and educational purposes.

Students' lives outside of school and college are increasingly media-rich and stimulating – which reflects the wider technological revolution taking place in society. As a result, education is becoming disconnected. The classroom is a sort of 'virtual reality'; a drab, technology-free zone that bears little relation to the increasingly technological reality of the students' real lives.

#### IV. ASSESSMENT 2.0

This paper suggests a way of modernizing assessment. The updated system will embrace the Internet and, more specifically, Web 2.0 – particularly the four "big ideas" described earlier. It seeks to bring the 21<sup>st</sup> century into the examination room. It attempts to do this not by 'upgrading' or 'patching' e-assessment (through simulations or machine marking or other 'fixes') – which can be considered Assessment 1.6 or 1.7 – but by using the same tools and techniques that students use at home and we use in the workplace.

The type of assessment activity best suited to the contemporary learner would exhibit some or all of the following characteristics.

**Authentic:** involving real-world knowledge and skills.

**Personalized:** tailored to the knowledge, skills and interests of each student.

**Negotiated:** agreed between the learner and the teacher.

**Engaging:** involving the personal interests of the student.

**Recognize existing skills:** willing to accredit the student's existing work.

**Deep:** assessing deep knowledge – not memorization.

**Problem oriented:** original tasks requiring genuine problem solving skills.

**Collaboratively produced:** produced in partnership with fellow students.

**Peer and self assessed:** involving self reflection and peer review.

**Tool supported:** encouraging the use of ICT.

**Personalized assessment** does not mandate individualized assessment (the setting of a unique task for each student). The teacher can continue to set the broad parameters of assessment activity; however, there may be flexibility in terms of time, place, contents, context and scope. At more advanced levels, learners may propose additional assessment criteria (to match their specific assessment activity) and there may also be an element of self- and peer-assessment permitted within the rubric.

The type of evidence that best fits this form of assessment would be:

**naturally occurring:** already in existence or generated out of personal interest

**multimedia:** existing in text, audio and video format

**digital:** such as e-mail, instant message logs, blog posts, wiki contributions, audio and video recordings

**distributed:** may be scattered across various sources (such as web sites, blogs, inbox, iPod).

Not all "assessment 2.0" tasks would embrace all of the above characteristics or media. But a modern assessment should seek to incorporate some of these characteristics and, certainly, permit the use of ICT.

For example, a "traditional" assessment task relating to History might ask students to describe the rise of Nazism in Germany in the 1930's. The evidence would be an essay, produced alone, under controlled conditions without reference to notes or other support materials. The equivalent "assessment 2.0" task would set the broad area of investigation (the rise of Nazism) but allow each student to choose a specific topic (such as the support given to the Nazi Party by US corporations). The assignment would be done collaboratively, in groups set by the teacher, with each member of the group choosing a specific sub-topic to research (such as the Nazis' use of IBM computers). The evidence would be in the form of a group blog, where each member of the team would post their findings (which would include hyperlinks to webpages, audio and video material) and the assessment would involve an element of self- and peer-assessment (along with teacher assessment). Unlike the essay, the blog may not require students to make any conclusions beyond reporting their findings via the blog, on the basis that any conclusions reached by the average 16 year old about such a complex period in history are likely to be superficial. This would contrast with the traditional approach, which would require a structured essay that is effectively an academic paper written by a child – and one greatly inferior to that found on Wikipedia (unless it was copied from there).

Two characteristics of modern assessment are particularly challenging for institutions: peer- and self-assessment, and assessing collaborative work.

##### A. Peer- and self-assessment

An important characteristic of a modern assessment is the use of self- and peer-assessment. These would not replace teacher assessment but may be an aspect of the overall marking rubric. In spite of being valid forms of assessment, self- and peer-assessment are rarely used in

tertiary education (except in formative assessment or very low stakes assessment) and almost never used in the school sector. It remains a challenge to bolster confidence in these forms of assessment and to find ways of applying them in summative assessment.

### B. Rubrics for collaboration

Web 2.0 provides a raft of tools to facilitate group work such as mailing lists, online forums, blogs, wikis and virtual worlds. However, the theory has not caught up with the practice. Existing marking schemes to assess contributions and interactions in these environments are vague or simplistic, often employing crude metrics such as the number of posts to a forum or subjective criteria such as “the quality of contributions”. These problems are multiplied in richer interactive environments, such as *Second Life*, where there has been little research into the measurement of effective contributions.

## V. HOW WEB 2.0 CAN BE USED FOR ASSESSMENT

Assessment is about evidence generation. The diagram below illustrates how evidence is traditionally produced.



Figure 1. Traditional evidence production

Evidence has to be discovered (when it already exists) or created (when it does not). The resulting information has to be captured and organized. And, once it is coherent, the evidence has to be assessed.

It is straight-forward to relate this model to Web 2.0. For example, a contemporary web-based e-mail system (such as Google Mail) can be used as a repository of every e-mail message you ever send or receive – which could be an Aladdin’s Cave of assessment evidence.

Downes (2006) describes the combination of Web 2.0 services for learning as “personal learning environments” (PLEs), arguing that the PLE is a “recognition that one-size-fits-all approach of [VLEs] will not be sufficient to meet the varied needs of students”. Assessment 2.0 posits Web 2.0 as a **personal assessment environment** in recognition that the one-size-fits-all approach of e-assessment systems will not be sufficient to meet the varied needs (and interests) of candidates.

Given that Web 2.0 is Life 1.0 for most students, it is an easy fit for most young people. They are already using Web 2.0 services as part of their everyday lives. Recognizing their MySpace page or their YouTube channel seems only ‘fair’ to them. And in doing so, it would re-

duce the perceived chasm between education and ‘real life’. It would also provide an incentive to learners: instead of artificial tasks involving “ancient” practices (such as hand-writing or using the library), assessment could provide real challenges using real tools – the same tools that they currently use outside of class and will use in the workplace.

Web 2.0 is inherently collaborative and the antithesis of Assessment 1.0’s obsession with individuality – and collaboration is a skill much sought after by employers. Web 2.0 services are also inexpensive (or free), easy to maintain (since they are maintained by someone else), and very scaleable (in fact, the more users the better). The alternatives (dedicated e-testing systems and e-portfolios) are proprietary, expensive, difficult to maintain and quickly become out-of-date. Assessment 2.0 is an approach, not a toolset. Facebook, Blogger *et al* may come and go but social networks and blogs are part of contemporary culture and will be around for the foreseeable future. Assessment 2.0 describes a type of task and an approach to that task; the choice of tools is up to each student. Maintenance and obsolescence are moot points.

There are drawbacks. Older students (our digital immigrants) aren’t using Web 2.0 services – or, at least, not routinely. They don’t have MySpace pages or YouTube videos to be plundered for accreditation of prior learning. And they may lack some key Web 2.0 skills (such as search skills) and attitudes (such as a willingness to share). Our digital natives themselves may not want to mix their public and private lives in the way suggested in this paper. They may want a clear dividing line between their academic lives and their personal lives, and may resist attempts to mix the two... although the barriers between social, educational and professional lives will have to be overcome if lifelong learning is to become a reality.

Assessment 2.0 also poses challenges for teachers – who are often the epitome of the digital immigrant. Not only might they lack the IT skills needed to understand Web 2.0 services but they may lack the knowledge and experience required to appraise students’ work produced using these tools.

Teachers also lack the rubrics required to assess Web 2.0 skills. Rubrics are required to address self- and peer-assessment, and collaboration. Group work is notoriously difficult to assess – so difficult that most awarding bodies prohibit it from high stakes assessment. Yet, it is at the core of Web 2.0 and a crucial skill for the workplace. New media presents new environments for students – environments such as *Second Life* and *The Sims*... and new challenges for awarding bodies to develop marking schemes to appraise student activity in these domains. Authentication is another problem for awarding bodies in the world of Assessment 2.0, with the myriad sources of digital evidence and collaborative inputs making it difficult to isolate and prove ownership of an individual piece of work. Self- and peer-assessment have proven to be valid and authentic forms of assessment but neither type of assessment has been widely used in schools or universities.

## VI. THE FUTURE

It’s impossible to predict the future with certainty. But there are certain themes that emerge when you review the literature relating to the future of education and technol-

ogy. With regard to education, there is a consensus about the following:

- greater focus on education as a key differentiator between countries in the global economy
- growth in learning at all stages in your life (the “fifty year degree program”)
- emergence of new skills to better fit the networked information economy
- greater role for e-learning (including mobile learning)
- move towards personalized learning (and, by corollary, personalized assessment)
- greater recognition of informal learning.

In tandem with these educational developments, the next decade will see two major technological developments: the emergence of ‘ubiquitous computing’ and ‘Web 3.0’. Ubiquitous computing describes a state of pervasive computing where digital devices are embedded into everyday life to such an extent that we are unaware of their existence. Web 3.0 will further develop the “big ideas” behind Web 2.0, particularly enhancing the intuitive and collaborative aspects of the Web. The cumulative effects of these trends will be an explosion of digitization, communication and collaboration.

If you combine these developments, you see a digitally rich environment where learning will take place in multiple locations (at school, at home, on the bus) at a time to suit the learner; where learning is personalized – in fact, a world where the distinction between learning and living is blurred and assessment evidence occurs naturally as part of the student’s everyday endeavors. It will be a world where Bill Gates’ vision of “information at your fingertips” will become a reality and where examinations that assess memorization (as most do today) will become untenable.

It is hard to imagine, in this future, any place for the ‘traditional’ VLE or CAA system. At best, posterity may view them as necessary stepping stones to the future; at worst, the final staging post of the educational establishment seeking to control learning. Current attempts at fixing them using expensive simulations or machine marking seem doomed to failure – a black hole through which money may be poured but from which nothing (lasting) emerges.

## VII. CONCLUSION

Assessment is often accused of preventing change. Critics claim high stakes assessment dictates what is taught and stifles innovation. So, if education is to change, that change has to be led by the assessment system – and contemporary e-assessment systems might not be the best way of doing that. Assessment 1.5 (or 1.6 or 1.7) can’t win the feature war with Web 2.0. What is state-of-the-art in an e-assessment system today (say, the inclusion of video in an online test) is state-of-last-year on the Web. They can never keep-up. And they will always feel unnatural to students.

One of the ways assessment can evolve is to adopt some of the characteristics of ‘Assessment 2.0’. That

means embracing Web 2.0 and the digital environments that students inhabit. It means moving from the analogue world of the past to the digital world of the future. Doing so would present a challenge to teachers and awarding bodies. Teachers would have to up-skill to better understand Web 2.0 and appreciate the world of the digital native. Awarding bodies would have to face the challenge of creating rubrics for assessing difficult to measure skills, such as collaboration, and confront issues such as plagiarism and peer-assessment. Both teachers and awarding bodies would have to embrace digital evidence in all of its forms and set more authentic tasks that genuinely challenge (and engage) students.

This paper is long on criticism and short on solution. But there is something wrong with the assessment system. It does need to change. The defenders of the *status quo* claim that many technologies promised to revolutionize education – but came and went without much of an impact on teaching and learning. Proponents of change point out that technologies such as TV and radio *did* revolutionize learning – just not the learning that happened to be taught and assessed in schools. To the advocates of change, similarly ignoring the information revolution will be impossible. Continued resistance will further marginalize education until change is forced on us.

## REFERENCES

- [1] Anderson, P (2007). What is Web 2.0? *JISC Technology and Standards Watch*. pp. 14-26.
- [2] Cousin, G (2003). *Learning from Cyberspace*. JISC.
- [3] DfES. Johnson Welcomes Rising Achievement. Retrieved 22 April 2007. [http://www.dfes.gov.uk/pns/DisplayPN.cgi?pn\\_id=2006\\_0119](http://www.dfes.gov.uk/pns/DisplayPN.cgi?pn_id=2006_0119)
- [4] Downes, S (2007). Learning Networks in Practice. *Emerging Technologies for Learning*, Volume 2, Chapter 2 (2007). Becta.
- [5] JISC. Case Studies of Effective and Innovative Practice. June 2006
- [6] Oblinger, D (2003). The Millenials. *Educause*. August 2003.
- [7] Owen, Grant, Sayers and Facer (2006). Social Software and Social Learning. *Futurelab, Opening Education Series*.
- [8] Pellegrino, JW (1999). The Evolution of Educational Assessment. *William Angoff Memorial Lecture Series* (ETS).
- [9] Prensky, M (2001). Digital Natives, Digital Immigrants. *On the Horizon* (NCB University Press, Vol. 9 No. 5, October 2001).
- [10] QCA. GCSEs and A level: the experiences of teachers, students, parents and the general public.
- [11] QCA. Evaluation of the 2006 pilot of the key stage 3 ICT test. September 2006. *QCA Assessment Research Team*.
- [12] Thompson, J (2006). Is Education 1.0 Ready for Web 2.0 Students? Nova Southeastern University.
- [13] Windham, C (2005). The Student’s Perspective: Educating the Net Generation. Chapter 5. *Educause E-book*.

## AUTHORS

**Bobby Elliott** is with the Scottish Qualifications Authority, Glasgow, G2 8DQ UK (e-mail: [bobby.elliott@sqa.org.uk](mailto:bobby.elliott@sqa.org.uk)).

This article was modified from a presentation at the 4th Competence Open Workshop in Madrid, Spain, April 2008. Manuscript received 14 May, 2008. Published as submitted by the author.